

RETAIL STORE SALES ANALYSIS

Fazil Mohammed



AGENDA

Introduction

Problem

Data Source

Data Cleaning & Manipulation

Analysis

Insights

INTRODUCTION

In this presentation, we will take a look at our retail store sales dashboard. This dashboard provides us with a quick overview of our sales performance, including units sold, revenue generated, and customer traffic.

PROBLEM

The leadership is interested in viewing the metrics from both an operations and marketing perspective. Management also intends to expand the business and is interested in seeking guidance into areas that are performing well so they can keep a clear focus on what's working.

Here are some requirements of the executives:

1. Show the revenue data for the year 2011.
2. Which are the top 10 countries generating the highest revenue?
3. Show the top 10 customers by revenue.
4. which regions have the greatest demand for our products.

DATA SOURCE

- The dataset was available from the Daqing Chen, Sai Liang Sain, and Kun Guo, Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining, Journal of Database
- It includes the details of retail sales in the year 2010 and 2011.
- The dataset contains 8 columns and 541910 rows of data running from the VoiceNo., StockCode, Description, Quantity, InvoiceDate, Unitprice, CustomerID and Country.

DATA CLEANING & MANIPULATION

- The dataset contains some returns to the store which are provided in negative quantities and there are unit prices which were input in error.
- Changed the negative quantities into 1 unit.
- Changed the negative unit prices into \$0.
- Null values were present in the Description field and filled with NA.
- Date and time was there in a single cell and then splitted into two separate columns with the help of Text to Column feature.
- A new field named Sales added to the dataset and filled with values by multiplying the UnitPrice and Quantity columns.

SUMMARY OF ANALYSIS

- Data was downloaded and stored using the appropriate file naming convention.
- Tools employed for analysis were Excel.
- The dataset was opened with Excel and the columns with null values were filled with NA, about 1455 in number.
- The data was sorted and filtered to make sure there was no misspelt or misformatted value under the columns.
- The date column splitted into two columns, one is for time.
- A new column name Sales added to the given dataset.
- The table then saved and uploaded to the Tableau Public.
- Analysis was noted and created an interactive dashboard with the help of Tableau Public

ANALYSIS

1. Revenue data for the year 2011.

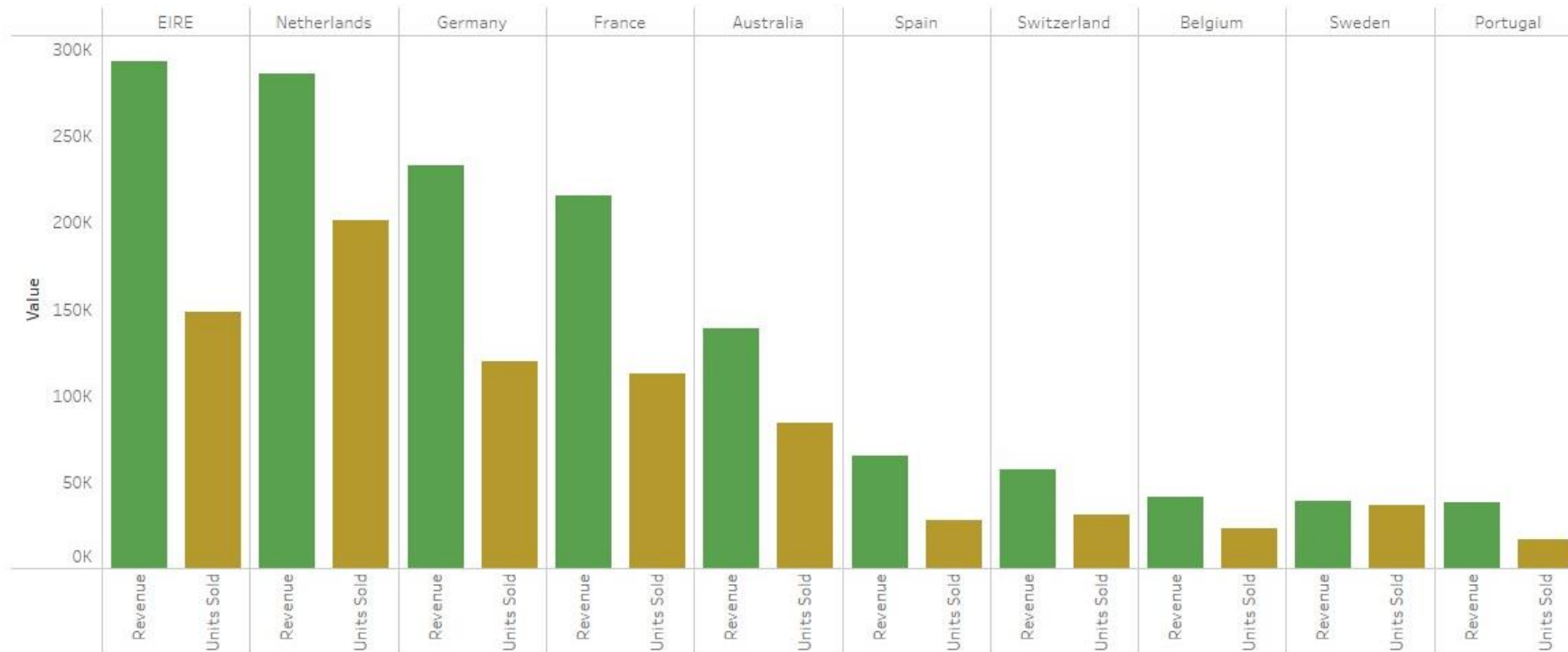
In 2011, store has generated the revenue around 1M. And most of the sales occurred in the month of Nov.



ANALYSIS

2. Top 10 countries generating the highest revenue

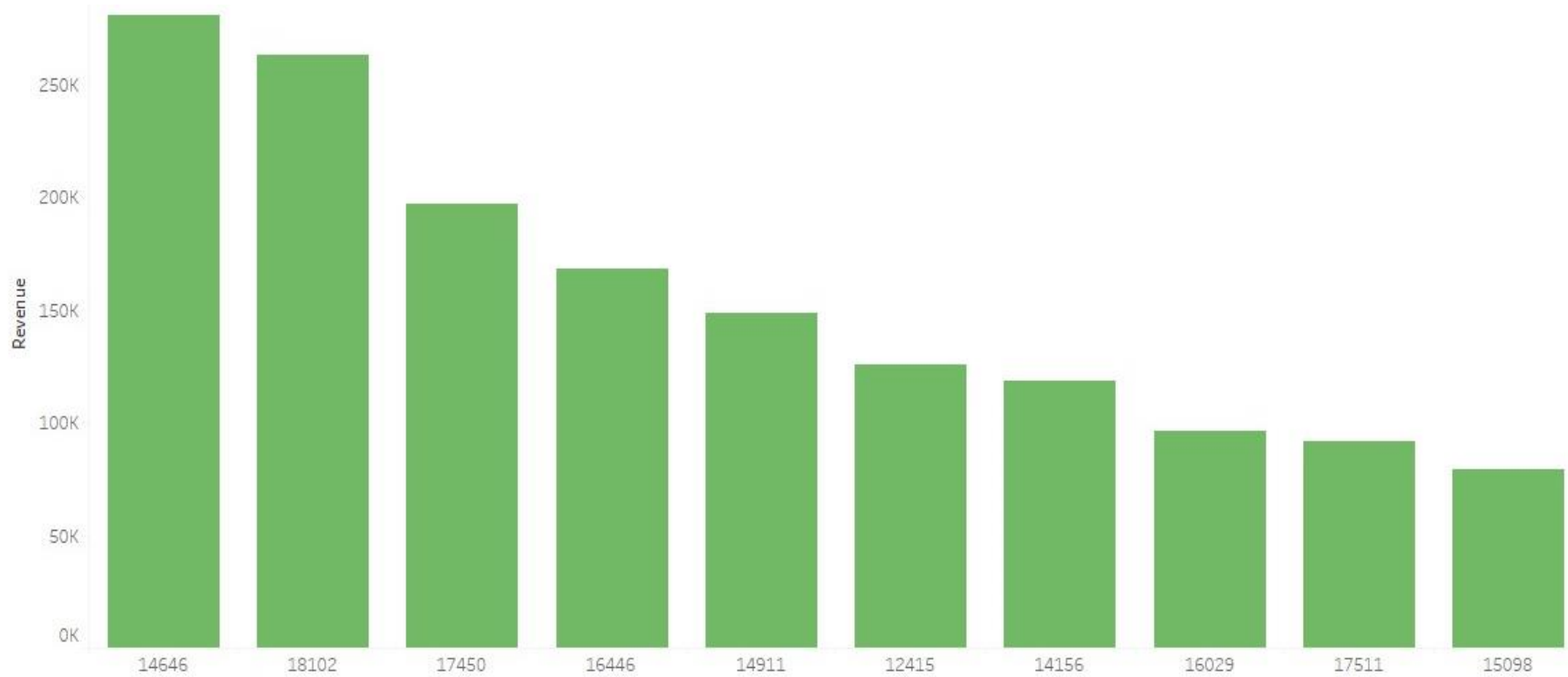
The Ireland is the top-selling country, generating \$290,000 in revenue. Netherlands and Germany follow closely behind, with \$280,000 and \$230,000 in revenue, respectively.



ANALYSIS

3. Top 10 customers by revenue

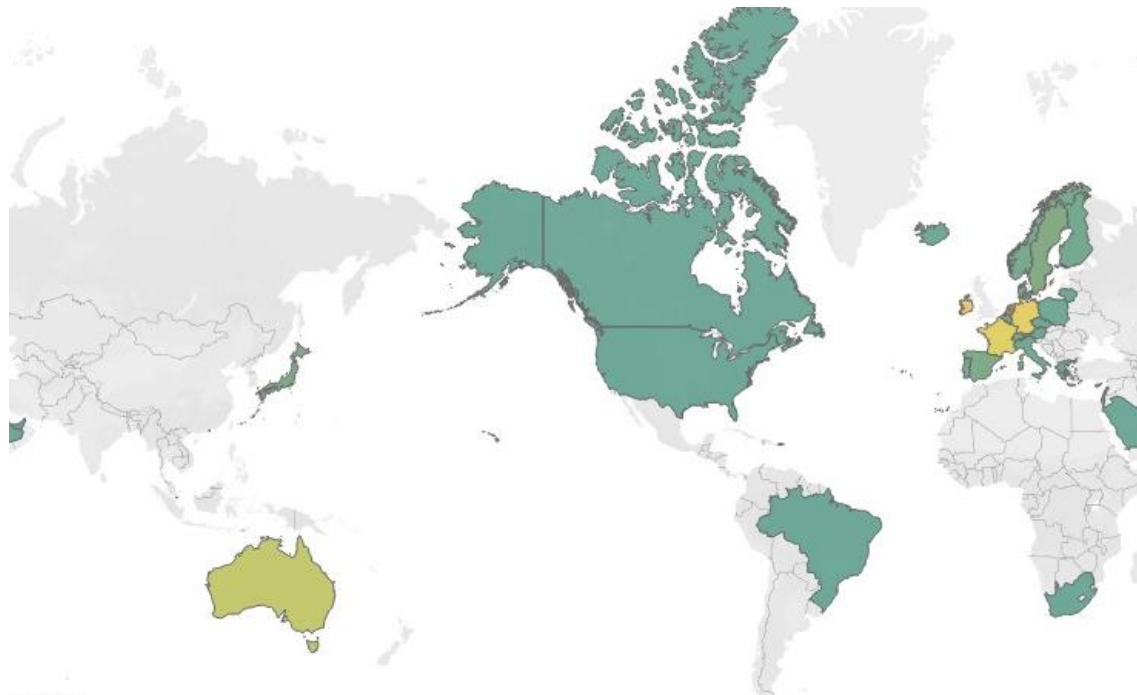
Customer 14646 is the top spender, generating \$280430 in revenue. Customer 18102 came in second, generating \$262876 in revenue. And Customer 17450 ranked third, generating \$196727 in revenue.



ANALYSIS

4. Regions with greatest demand for our products

Europe appears to be the top region based on revenue, with France, Germany, Netherlands, Spain, Switzerland, and Portugal contributing significantly. North America likely the second-highest contributor based on revenue, though the data only shows one country, the United States.



INSIGHTS

- Cultural factors or product preferences could also play a role. Certain products might be more popular in specific countries, leading to higher sales in those regions.
- Economic factors like disposable income or exchange rates could also influence purchasing behavior. Customers in countries with higher disposable income or favorable exchange rates might be more likely to spend more.
- There is a noticeable gap in revenue between the top 3 customers and the remaining customers on the list.
- It would be beneficial to understand the factors contributing to the high revenue generated by the top customers. This could involve analyzing their purchase history, product preferences, or any loyalty programs they might be participating in.
- The data suggests some presence in South America (Brazil) and Asia (Singapore), but it's difficult to assess their relative contribution compared to Europe and North America due to missing data.

THANK YOU